

BMJ Open How different countries addressed the sudden growth of e-cigarettes in an online tobacco control community

Kar-Hai Chu, Thomas W Valente

To cite: Chu K-H, Valente TW. How different countries addressed the sudden growth of e-cigarettes in an online tobacco control community. *BMJ Open* 2015;**5**:e007654. doi:10.1136/bmjopen-2015-007654

► Prepublication history for this paper is available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2015-007654>).

Received 13 January 2015
Revised 15 April 2015
Accepted 29 April 2015



CrossMark

Department of Preventive Medicine, University of Southern California, Los Angeles, California, USA

Correspondence to
Dr Kar-Hai Chu;
karhaich@usc.edu

ABSTRACT

Objective: The sudden growth of e-cigarettes over the last decade has forced advocates and critics scrambling to bolster support for their respective sides. Bridging the divide in geographic barriers, social networking sites were an ideal meeting place for international activist communities, affording them the ability to organise events and discuss new topics in real time. This study examines how e-cigarettes are addressed in GLOBALink, an online international tobacco control community. We seek to discover if the pattern of activity in e-cigarette discussions changes over time. We are also interested in understanding the characteristics of sentiment toward e-cigarettes in discussion topics between countries with different network characteristics.

Design: Network analysis to explore the relationships between members from different countries, and sentiment analysis of messages and threads to identify patterns of how different countries address e-cigarette topics.

Setting: GLOBALink, an online international tobacco control community.

Participants: Network analysis based on GLOBALink members from 37 different countries. Sentiment analysis based on 853 posted messages, with over 1.4 million words.

Outcome measures: Network centrality measures in country interaction data, including degree, closeness and betweenness. Sentiment scores for each message, and differences between country scores.

Results: The network analysis found a core/periphery structure where central countries focused on active positive discussions pertaining to e-cigarettes, while isolated and peripheral countries posted negative topics without many responses. A qualitative examination of message topics suggests that general subjects elicit more interactions than those that are context specific.

Conclusions: E-cigarettes are a polarising topic that can be seen in how countries appear to discuss related topics with others who share the same opinions. More work is needed to help communities stay informed of current research, and diffuse objective information. Network and sentiment analyses offer a strong combination of methodologies that can help support such efforts.

Strengths and limitations of this study

- The study of e-cigarettes has typically focused on awareness and the potential for benefits or harm. This paper takes a different approach, and examines an international community of tobacco control advocates, at a time when e-cigarettes were still relatively unknown.
- We apply modern network and sentiment analysis methods to provide a rich understanding of the attitudes and relationships in how a diverse community faced the challenge of an unexpected development in their efforts to support global tobacco control. Our findings can help identify—even within a group of like-minded advocates—where gaps in communication, or differences in opinions, might exist and how they might be resolved.
- One concern in using a dictionary of predefined sentiment scores is the issue of context. General scores are typically derived from large corpuses, but topic-specific studies might be sensitive to how certain terms are used. Applying a more customised dictionary with improved semantic algorithms could help generate more accurate sentiment scores.

INTRODUCTION

Information and communication technologies have become part of our everyday lives. The emergence of Web 2.0 technologies affords the casual user a more active and participatory role in all aspects of the online world.¹ This is especially true of social networking sites (SNS), where niche communities have developed rapidly.² SNS have demonstrated that people can and do form and sustain successful relationships online. Internet media has enabled the growth of activist networks and created a rich database of interactions regarding various policies, social movements, and other causes. For example, sites such as Facebook and Twitter have been central in helping spread the Occupy Wall Street³ and Arab Spring⁴ movements. These actions were well publicised

and serve as tangible examples of internet activism. The involved groups were able to improve coordination of activities in a large-scale setting, recruit high-profile celebrities to their causes, and record and display information for the public to see. Tools within SNS afforded members the ability to act within certain roles, whether to initiate events, spread information or coordinate efforts. The online setting also helped break through geographic barriers and support international communication. It is now relatively easy for activists to conduct and coordinate events around the world.

In a parallel narrative, e-cigarettes have witnessed a similarly rapid growth. E-cigarettes reproduce the effect of tobacco cigarettes by heating a liquid to create a nicotine vapour, which can then be inhaled. E-cigarettes sales are estimated to be over US\$1 billion in 2013, and are expected to continue growing. Major tobacco companies have recognised their value, and have been introducing their own e-cigarette brands (eg, Reynolds launching VUSE) or acquiring smaller e-cigarette companies (eg, Lorillard purchasing Blu).

The push by tobacco control advocates against e-cigarettes has been divided. This is likely due to the possibility that it is a viable method to quit smoking,⁵ as findings from several recent randomised controlled trials have suggested.^{6 7} Research on e-cigarettes has otherwise focused on awareness,^{8 9} harms¹⁰ and benefits.¹¹ There have also been grassroots movements that support e-cigarettes, using social media to organise events and rallies. For example, IMPROOF, a group of e-cigarette supporters (<http://www.improofmovement.com>), builds online activism through Twitter (use of the hashtag #IMPROOF) and other social media websites to spread information, recruit members and plan gatherings.

By contrast, we examine a community of tobacco control advocates, made up of administrators, researchers, policymakers and others who are fighting tobacco globally. Our study investigates GLOBALink, an online tobacco control community that has been active for over 20 years. In this study, we examine the network of GLOBALink members in the context of their shared interactions in an online discussion forum. We will also examine the content of their messages and perform a sentiment analysis to gauge whether different countries have positive or negative opinions about e-cigarettes.

BACKGROUND

Social network analysis has been applied to identify actor roles in various situations, for example, in the diffusion of innovations,¹² online conversations,¹³ organisational structures¹⁴ and so on. We study the interactions in GLOBALink's discussion forum. Asynchronous discussion forums have been popular virtual spaces that allow people to congregate and discuss topics of shared interest. Several studies^{15 16} have examined growth patterns and membership adoption in modern discussion-based

communities, and compare them with diffusion models. Content analysis¹⁷ has also been applied to data derived from discussion forums.

In 1992, the Switzerland-based Union for International Cancer Control (UICC) took over coordination of a small US-based tobacco control network and formed GLOBALink, an online network of tobacco control professionals. Over the next two decades, GLOBALink grew into a large network with members from across the world dedicated to controlling tobacco use. GLOBALink's homepage contains news bulletins, electronic conferences, live interactive chat and full-text databases (including news, legislation, directories). Other studies on GLOBALink have found the community to be a strong influence on global tobacco control. Wipfli and colleagues¹⁸ found that the likelihood of ratifying the WHO Framework Convention on Tobacco Control by a country was three times as likely when that country was exposed to other members of ratifying countries via membership in GLOBALink. Chu and others¹⁹ presented a longitudinal exploratory study with dynamic visualisations that helped understand the evolution of communities in the GLOBALink network.

We propose several research questions (RQ, below) for the network and content analyses of GLOBALink messages regarding e-cigarettes:

- RQ1: What is the pattern of activity in e-cigarette discussions over time?
- RQ2: What is the pattern of sentiment in discussion topics between countries with different network characteristics?

GLOBALink's network of countries can exhibit many possible properties. In Chu *et al's*¹⁹ study, the referral network (through which new members can only join after being referred by two existing members) showed how clusters of countries were formed based on geographic location, political affiliations and colonial history. We expect different network characteristics in the discussion network, and would like to see if the network properties correlate with the types of discussion topics being posted. In this manner, we can start identifying which countries are discussing what topics, and how cross-cluster conversations might occur.

METHODOLOGY

In this study, we examine data from GLOBALink and begin with an exploratory network analysis, followed by a more thorough content analysis.

Data

The data from GLOBALink was received as a comma-separated values flat file and loaded into a MySQL database. Express permission and support was given by the UICC, the organisation that hosted GLOBALink during the time period for which we analysed the data. Use of the data in this study has also been reviewed by the Institutional Review Board of the University of Southern

California and determined to be exempt. Relevant message data included the identifier (ID) of each message, the ID of the discussion thread, the country of the user who posted the message, the subforum where the message was posted, and the date of posting. All user data are kept private, as we aggregate the message subjects to the country level, effectively removing information about the individual who posted the message. Additionally, no user-posted text is directly quoted in this manuscript. The data cover all messages from November 2004 to May 2012.

Exploratory network analysis

We began with a network analysis using the discussion forum data. We performed a search of all message headers and bodies in the MySQL database that included any of the following terms: 'e-cig', 'e cig', 'electronic-cig' and 'electronic cig'. After finding 900 possible matches, we randomly sampled 200 messages to determine the accuracy of our search terms. We manually removed irrelevant messages that were captured due to the relaxed nature of the search algorithm and non-English postings. Conversely, we also used the results to help find additional terms that could be related (eg, 'electric cig' was found in many results, and added for new searches). Several more iterations were run, repeating the same sample cleaning process. After we completed the additions and removals, we had a final sample size of 853 messages, posted by members in 37 countries, from July 2005 to April 2012.

Each posted message is part of a discussion thread, where any number of other members can respond. By linking together all members in the same discussion thread, we constructed a network of countries based on their shared presence in the threads. The network data are dyads in the form of 'country-country' relationships. Network visualisations are then created from these dyadic relationships, using the Gephi software package (<https://gephi.org>).

We next follow the network of countries by 'unpacking' all its ties. As a tie represents discussion threads that are shared by any two countries, we can view the network with each discussion thread exposed as additional nodes. We transform the 'country-country' data into 'country-thread-country' data, and then break the triad into two 'country-thread' dyads. This is called a bipartite, or 2-mode network (see refs. 20 and 21 for explanations on working with 2-mode data). This 2-mode data help us visualise the relationships between countries or discussion threads, and to identify significant structural properties.

Sentiment analysis

The content analysis is conducted within the MySQL database with custom scripts. Using the 853 messages found in the network analysis, we perform a sentiment analysis of the messages to identify the opinions of e-cigarettes in the community. To determine if a message is positive or negative, we use a simple bag-of-words

model²² of classifying the terms found in each message. The dictionary of words comes from the Multi-Perspective Question Answering (MPQA) Subjectivity Lexicon (<http://mpqa.cs.pitt.edu>), which identifies 6451 words as positive or negative, with an additional strong or weak quantifier. From the 853 messages concerning e-cigarettes, there are over 1.4 million words in the text. For each message, we compare every word and attempt to match it against the terms in the MPQA dictionary. If the word is not found, we also apply a stemming algorithm to see if the root word is available. For example, *afflicted* is not found in the sentiment list, but we can stem the word to *afflict*, which is found in the list. If the word, or its stemmed root, is found, we apply a score to the message:

- ▶ Strong, positive = +2
- ▶ Weak, positive = +1
- ▶ Weak, negative = -1
- ▶ Strong, negative = -2

Because messages can be very different in length, the raw scores are inadequate for comparison. In addition to the raw scores, we also normalise the scores to control for message size.

We conduct several tests to discover how sentiment might connect with different components in the network. First, we examine how sentiment scores for e-cigarettes compare against topics not related to e-cigarettes using an independent samples t test. We also use results of the network analysis to find any metrics that might connect country interactions with the sentiment scores.

RESULTS

Our final dataset consists of 853 messages posted by members in 37 countries, from July 2005 to April 2012. The number of posts over time can be seen in figure 1.

Network analysis

Figure 2 depicts how countries (represented as nodes, or vertices) are linked to each other. A tie connects two countries if they coparticipate in at least one discussion thread (ie, both postmessages in a single thread). The strength of the tie—depicted visually by the thickness of the line—is greater if the two countries share a presence in many discussion threads. The size of the node represents degree centrality, or the number of other countries a node is connected to.

In the 2-mode network (figure 3), red nodes represent countries and blue nodes represent discussion threads. Each tie now links a country with discussion threads that have been posted by members of that country. Node sizes for each country (ie, red nodes) are reset so they are all the same, but we adjust the discussion threads' (ie, blue nodes) size based on their betweenness centrality. Betweenness is a network measure that indicates how frequently a node lies on the shortest path between all pairs of nodes; the more

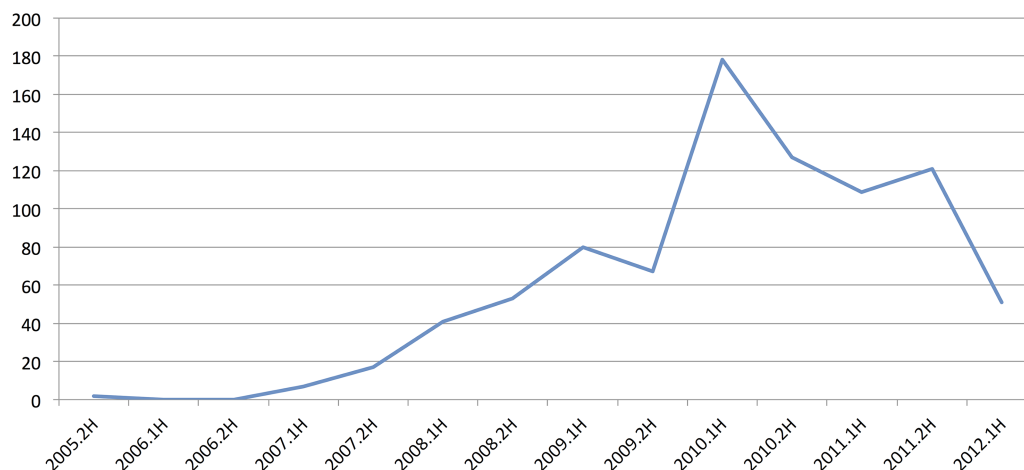


Figure 1 Number of messages posted about e-cigarettes over time.

number of shortest paths it resides in, the higher the betweenness value.²³ In this context, the larger blue nodes represent discussion threads that directly link many countries together when they otherwise might not be connected. We also calculate closeness centrality (not represented visually), which measures the distance any node is to all other nodes. Generally, core nodes will have higher closeness, as they have shorter paths to all other nodes than those on the periphery.

With the 2-mode network, we now have a clear picture of the pattern of interactions in the GLOBALink forums. We have labelled several nodes of interest and have identified them. First, we include the top five countries as determined by degree centrality (ie, number of discussion threads they are present in), which are the same five we had visually found in the country network's core cluster. Next, we label the top five discussion thread IDs, as determined by their betweenness centrality:

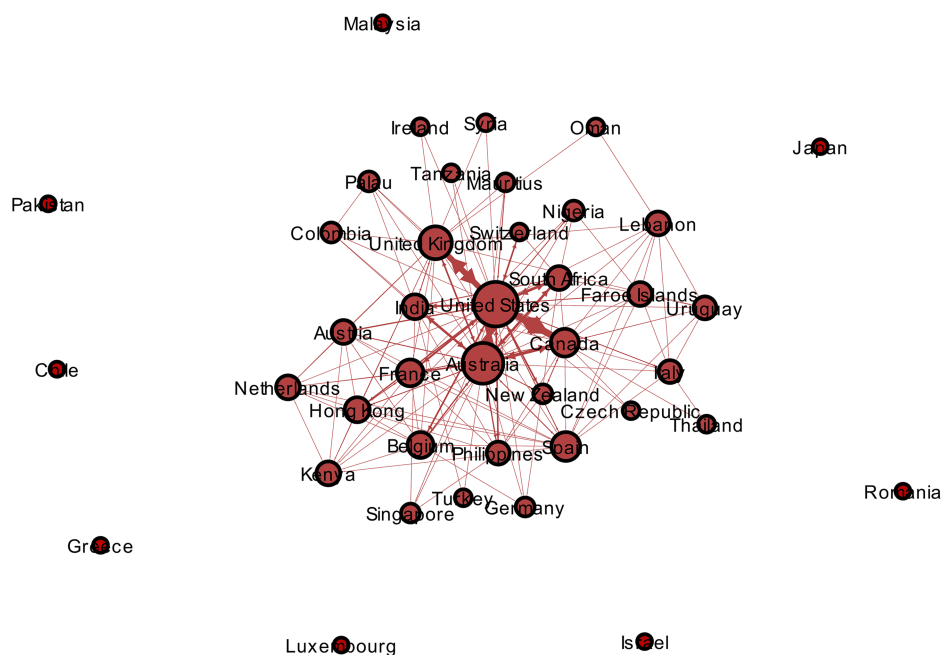
8324, 6, 13 022, 6467 and 9236. These threads serve to mediate discussions between many pairs of countries. Last, we collect the thread IDs for the discussions that are connected to the isolates (not labelled).

Sentiment analysis

Table 1 gives a general description of the sentiment scores for all the messages. Figure 4 shows the pattern of sentiment in each message over time.

To see how e-cigarettes compared with other topics in GLOBALink, an independent samples t test was conducted to compare the sentiment scores for the e-cigarette messages against all other messages in the same time period (July 2005–April 2012). There was a significant difference in the scores for e-cigarette messages ($M=0.0103$, $SD=0.0244$) and all other messages ($M=0.0144$, $SD=0.0294$); $t(41\,695)=-3.87$, $p<0.001$,

Figure 2 GLOBALink network of country-country interactions.



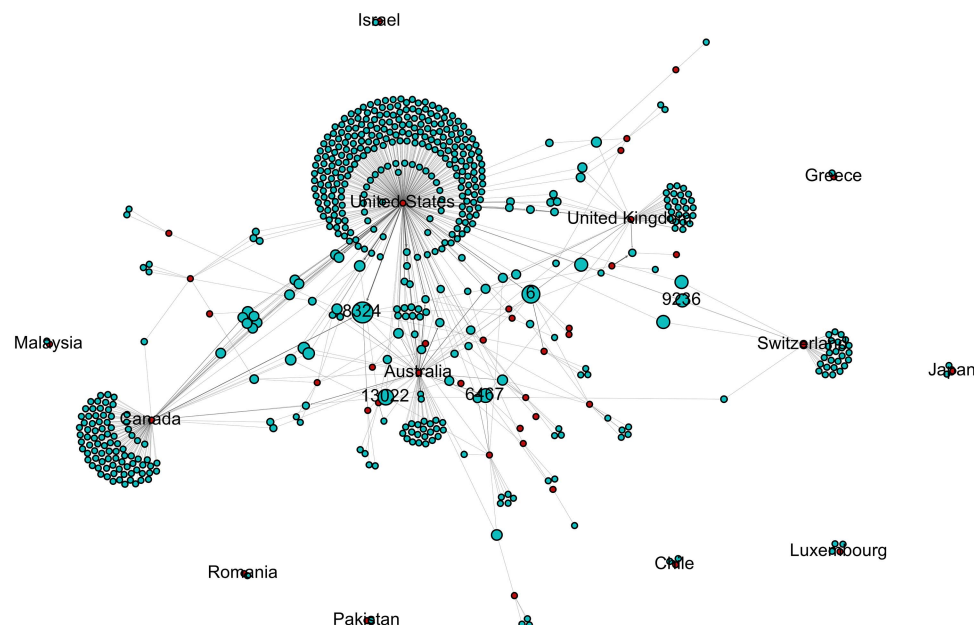


Figure 3 GLOBALink 2-mode network of country-thread interactions.

indicating that e-cigarette postings were significantly more negative.

A post hoc simple linear regression was conducted to examine if the difference in sentiment between e-cigarettes and other topics could be predicted by closeness centrality. The results were significant, $F(1,32) = 8.67$, $p < 0.01$, and accounted for 18.86% (adjusted R^2) of the explained variability. The regression equation was: predicted difference = $0.029 - 0.026 \times (\text{closeness centrality})$.

DISCUSSION

The exploratory network analysis provided data that helped inform the later content analysis. We can make several observations based on the country-country network graph (figure 2). The network shows a core/periphery structure, with several nodes in a closely connected dense centre surrounded by more loosely connected nodes at the outskirts. We can clearly see the high degree core countries, most notably the USA, Australia, Canada, Switzerland and the UK, indicating a very interactive group of countries that participated in many discussion threads together. At the other end of

the network, we also notice the eight disconnected nodes, or isolates: Pakistan, Malaysia, Japan, Greece, Chile, Romania, Luxembourg and Israel. Not having any ties with other countries means that the isolates, while posting discussion messages about e-cigarettes, were not involved in threads where other countries also participated. This difference would direct us to compare message topics to find out why certain topics attract more attention than others.

The second network graph (ie, the 2-mode network) provided data useful for examining the messages being posted. We use betweenness centrality in the visualisation (represented by node sizes) because it is a network measure that provides information about how important any given node is in connecting other nodes. Table 2 shows the topic headers and sentiment scores for the 12 threads with the highest betweenness, representing discussions that involved interactions between many countries. Table 3 includes the 12 threads that are connected to the isolate countries, that is, they did not foster any discussion. From an initial observation, it would appear there might be a trend showing that isolated threads tend to exhibit negative sentiment. All the high betweenness threads were positive, while 50% of the isolated threads were negative.

Even though we see a growth of e-cigarette message postings (figure 1), the overall trend in sentiment does not noticeably become more positive or negative (figure 4). Table 1 shows that there are more than twice as many positive than negative discussions. These descriptive statistics provide a simple answer to RQ1: that while more conversations are taking place about e-cigarettes as they become more popular, sentiment does not appear to change over the same period of time. To answer RQ2, we analysed the relationships between discussion sentiment and network characteristics.

Table 1 Description of messages and sentiment

Observations	853
Raw range of sentiment scores	-144 to 130
Mean sentiment score (SD)	11.34584 (30.05033)
Mean sentiment score normalised by word count (SD)	0.0103133 (0.0244054)
Messages with positive scores	528
Messages with negative scores	252
Messages neutral or unscored	73

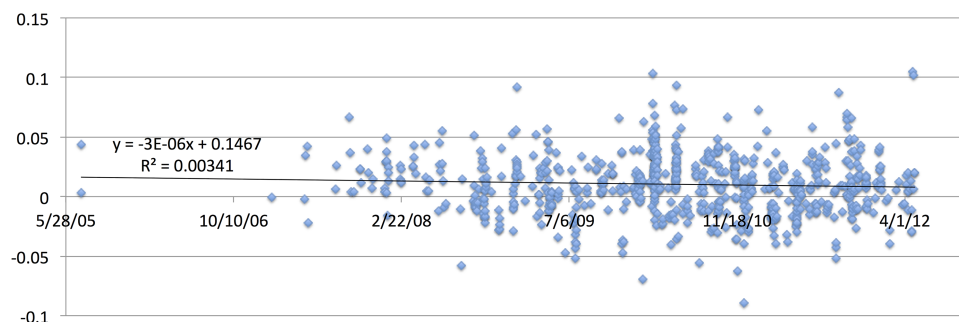


Figure 4 Sentiment of e-cigarette messages over time.

Post hoc tests

The results of the sentiment comparison test suggest that sentiment regarding e-cigarettes is generally more negative than other topics discussed in GLOBALink. We examined several other attributes of the same 853 messages and their related threads to identify potential network metrics that might help explain some of the difference. The top of table 4 consists of a list of the top five countries with the largest differences in their discussion sentiment between e-cigarette topics and all other topics. Each of the five countries is either an isolate in the e-cigarette discussion network (figure 2) or at the periphery of the connected group. By contrast, the bottom of table 4 includes the five central countries located at the core of the network. These five countries have very little difference in sentiment when comparing e-cigarette and all other topics; in fact, Switzerland and Canada actually have slightly more positive sentiment scores for e-cigarette topics. In the GLOBALink network, these results might be discouraging when viewed in the context of diffusing information and sharing ideas, but helps us to address RQ2. When looking for a pattern of how discussion topics vary between countries with different network characteristics, it would appear that the most active countries share

similar positive opinions on e-cigarettes and frequently interact with each other. At the outskirts of the network, countries that discuss e-cigarettes in a relatively negative manner are rarely in contact with other countries, if at all. This pattern of interactions creates homogeneous subnetworks where new ideas are not being exchanged, and countries with similar opinions only communicate with others that already share their beliefs. To test this, we conducted a simple linear regression analysis to examine if the difference in sentiment between e-cigarette topics and all other topics could be predicted by closeness centrality. The significance of the results suggests that the peripheral countries have significantly more negative e-cigarette discussions than core countries, confirming our visual findings for RQ2.

A more content-sensitive view of the topics and messages appeared to help explain some of the differences

Table 2 Top 12 threads based on betweenness, including information on topic and sentiment

Thread	Summarised message topic	Betweenness	Sentiment
8324	Asking for info	0.0415165	0.0148
6	General info	0.0341207	0.0214
13 022	General info	0.0266851	0.01872
6467	Industry packaging	0.0219485	0.032
9236	–	0.0203558	0.0038
10 772	Asking for info	0.0203558	0.01753
14 746	Health info	0.0203558	0.04325
15 596	Health info	0.0203558	0.00435
9381	–	0.0153913	0.0216
11 054	–	0.0153913	0.03243
11 960	Health info	0.0136741	0.02022
8504	Country bans of e-cigarettes	0.0131022	–0.00743

Table 3 12 isolated threads, including information on poster country, topic and sentiment score

Thread	Country	Summarised message topic	Sentiment
8475	Israel	General e-cigarette questions	0.0526
15 055	Greece	(topic)	–0.0135
11 011	Japan	Alternative smoke-free to e-cigarettes in Japan	–0.012
11 349	Japan	E-cigarette company in Japan	–0.0112
13 648	Luxembourg	European Union policy discussion	–0.0056
15 696	Luxembourg	(topic)	0.0201
15 695	Luxembourg	(topic)	0.0202
10 304	Chile	(topic)	0
10 306	Chile	(topic)	0.0034
9497	Pakistan	E-cigarettes in Pakistan, questions on harm reduction	–0.0476
1977	Romania	Questions on regulating e-cigarettes	0
11 837	Malaysia	Questions on 'stealth' e-cigarettes	–0.0273

Table 4 Ranks of 10 countries based on difference in sentiment scores between e-cigarette topics and all other topics

Country	Rank	E-cigarette sentiment	All other sentiment	Difference
Pakistan	1	−0.0476	0.00273953	−0.05033953
Malaysia	2	−0.0273	0.02150714	−0.04880714
Japan	3	−0.0116	0.03651304	−0.04811304
Colombia	4	−0.0333	0.01004545	−0.04334545
Ireland	5	0.005	0.03818923	−0.03318923
UK	16	0.00909773	0.02349269	−0.01439496
Australia	18	0.01133333	0.02331831	−0.01198498
USA	19	0.00845785	0.01930207	−0.01084422
Switzerland	22	0.01335641	0.00450547	+0.00885094
Canada	33	0.00868673	0.00804523	+0.0006415

Top five have the highest difference in sentiment scores. Bottom five are countries central in the network.

in other-country responses. Of the 12 topics with the highest betweenness (table 2), 9 were focused on e-cigarettes in general, while 3 were location specific. By contrast, in the 12 isolated topics (table 3), over 50% (7) were specific to either a location (eg, Japan, Argentina, Europe, Pakistan) or context (eg, US military). This could be due to each country having very different laws regarding tobacco control and e-cigarette use. These differences are less ‘open for debate,’ while information on e-cigarette usage, health and other location-neutral topics have more room for discussion.

It is also important to view the results of the analyses in a broader view, and understanding the difference in attitudes outside the network context. GLOBALink is a network of tobacco control advocates, but there is no clear consensus on their opinions toward e-cigarettes. While our results suggest that homogeneous subnetworks of countries are preventing cross-discussion of topics with opposing beliefs, it nonetheless reveals that e-cigarettes have managed to remain polarising over nearly 7 years of study. Current research still leaves the topic open for debate, as new studies are being conducted on the nature of the chemicals in the inhaled vapour, whether or not it is a viable method to help quit smoking, and awareness of the technology by different populations.

Limitations

One concern in using a dictionary of predefined sentiment scores is the issue of context. General scores are typically derived from large corpuses, but topic-specific studies might be sensitive to how certain terms are used. Applying a more customised dictionary with improved semantic algorithms could help generate more accurate sentiment scores. Similarly, sentiment dictionaries that are limited to English prevent us from expanding our study to include messages posted by GLOBALink’s non-English speaking members. Advanced machine-learning techniques have been explored to better study corpuses in different languages, but these are beyond the scope of our study. The authors are also involved in a parallel project using automated classification methods

to help understand the topics being discussed in GLOBALink. In conjunction with a refined sentiment analysis, this could reveal more intricate patterns in GLOBALink interactions.

CONCLUSION

The rapid growth of e-cigarettes demands attention from both supporters and detractors alike. As online activism becomes more accessible, each side has staked their claim in various SNS to promote their stance and gather supporters. GLOBALink is a large online network of tobacco control advocates with a history developed before the internet came of age. Having successfully fought for issues including the WHO Framework Convention on Tobacco Control, it serves as an ideal setting to study the impact of e-cigarette growth. We found that general sentiment regarding e-cigarettes tended to be more negative than other tobacco topics, but that it has not changed much over a 7-year period. Using a combination of network and sentiment analysis, we were able to find a core of highly central countries discussing e-cigarettes in a positive manner, while more isolated countries discussed negative aspects of the topic without much interaction. These findings suggest that e-cigarette discussions are not diffusing between countries with differing opinions, and demonstrate the need for common ground in discussions or for more moderate countries to try and build connections. Looking ahead, researchers must be conscious of the trend of social media websites slowly replacing communities such as GLOBALink. Recent investigations of e-cigarettes have taken place on these platforms, such as marketing on Twitter²⁴ and vape shops on Yelp.²⁵ In another example, Harris *et al*²⁶ followed the trend of Twitter posts related to the vote of e-cigarette regulation, finding that a majority of users were against it. The science behind e-cigarettes continues to grow, and objective, grounded and accurate information needs to be shared between any and all countries involved in the subject. Continued research into how e-cigarettes are discussed and viewed in online networks can help us work together to better understand the

impact that it could have, whether as a proven harm-reduction technology or a potentially dangerous device that introduces nicotine to non-smokers and teenagers.

Acknowledgements The authors would like to thank the Union for International Cancer Control for its assistance, and Stephanie Dyal for her help with the edits.

Contributors K-HC was the primary author, and contributed to the planning, study design, analyses and manuscript. TWV is the PI of the supporting grant, and contributed to the analyses and the manuscript.

Funding This research was supported by National Institute of Health grant number 1R01CA157577 (PI: Thomas W Valente) and training grant T32CA009492 from the National Cancer Institute/National Institutes of Health.

Competing interests None declared.

Ethics approval University of Southern California Institutional Review Board.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

REFERENCES

- Beer D, Burrows R. Sociology and, of and in Web 2.0: some initial considerations. *Sociol Res Online* 2007;12.
- boyd danah, Ellison N. Social network sites: definition, history, and scholarship. *J Comput-Mediat Commun* 2007;13:210–30.
- Caren N, Gaby S. Occupy online: Facebook and the spread of Occupy Wall Street. *Soc Sci Res Netw* 2011. <http://dx.doi.org/10.2139/ssrn.1943168>
- Hermida A, Lewis SC, Zamith R. Sourcing the arab spring: a case study of Andy Carvin's sources on twitter during the Tunisian and Egyptian revolutions. *J Comput-Mediat Commun* 2014;19:479–99.
- Polosa R, Caponnetto P, Morjaria JB, et al. Effect of an electronic nicotine delivery device (e-cigarette) on smoking reduction and cessation: a prospective 6-month pilot study. *BMC Public Health* 2011;11:786.
- Bullen C, Howe C, Laugesen M, et al. Electronic cigarettes for smoking cessation: a randomised controlled trial. *Lancet* 2013;382:1629–37.
- Caponnetto P, Campagna D, Cibella F, et al. Efficiency and Safety of an eElectronic cigAreTte (ECLAT) as tobacco cigarettes substitute: a prospective 12-month randomized control design study. *PLoS ONE* 2013;8:e66317.
- Emery SL, Vera L, Huang J, et al. Wanna know about vaping? Patterns of message exposure, seeking and sharing information about e-cigarettes across media platforms. *Tob Control* 2014;23 (suppl 3):iii17–25.
- Regan AK, Promoff G, Dube SR, et al. Electronic nicotine delivery systems: adult use and awareness of the “e-cigarette” in the USA. *Tob Control* 2013;22:19–23.
- Barrington-Trimis JL, Samet JM, McConnell R. Flavorings in electronic cigarettes: an unrecognized respiratory health hazard? *JAMA* 2014;312:2493–4 (cited 15 December 2014).
- Hajek P. Electronic cigarettes have a potential for huge public health benefit. *BMC Med* 2014;12:225 (cited 16 December 2014). <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4260378/> <http://dx.doi.org/10.1186/s12916-014-0225-z>
- Rogers E. *Diffusion of Innovations*. 5th edn. New York: Free Press, 2003.<http://www.amazon.de/exec/obidos/ASIN/0743222091>
- Suthers DD, Desiato C. *Exposing chat features through analysis of uptake between contributions*. Institute of Electrical and Electronics Engineers, Inc (IEEE); 2012.
- Cross R, Prusak L. The people who make organizations go—or stop. *Harv Bus Rev* 2002;80:104–12. 106 http://www.uniroma2.it/didattica/direzioneestrategie/deposito/cross_prusak.pdf
- Backstrom L, Huttenlocher D, Kleinberg J, et al. *Group formation in large social networks: membership, growth, and evolution*. ACM, 2006:44–54.
- Firth DR, Lawrence C, Clouse SF. Predicting internet-based online community size and time to peak membership using the bass model of new product growth. *Interdiscip J Inf Knowl Manag* 2006;1:1–12.
- Krippendorff K. *Content analysis: an introduction to its methodology*. SAGE Publications, Incorporated, 2012. <http://www.uk.sagepub.com/booksProdDesc.nav?prodId=Book234903>
- Wipfli H, Fujimoto K, Valente TW. Global tobacco control diffusion: the case of the framework convention on tobacco control. *Am J Public Health* 2010;100:1260.
- Chu KH, Wipfli H, Valente TW. Using visualizations to explore network dynamics. *J Soc Struct* 2013;14:pii: 4. <http://www.cmu.edu/joss/content/articles/volume14/ChuWipfliValente.pdf>
- Borgatti S. *2-mode concepts in social network analysis*. *Encyclopedia of complexity and systems*. Berlin: Springer, 2009:8279–91. <http://steveborgatti.com/papers/2modeconcepts.pdf>
- Valente TW. *Social networks and health*. Oxford: Oxford University Press, 2010. <http://www.amazon.com/Social-Networks-Health-Methods-Applications/dp/0195301013>
- Salton G, McGill MJ. *Introduction to modern information retrieval*. McGraw-Hill, 1983:470. <http://books.google.com/books?id=7f5TAAAMAAJ>
- Freeman LC. Centrality in social networks conceptual clarification. *Soc Netw* 1979;1:215–39.
- Huang J, Kornfield R, Szczypka G, et al. A cross-sectional examination of marketing of electronic cigarettes on Twitter. *Tob Control* 2014;23(suppl 3):iii26–30.
- Sussman S, Garcia R, Cruz TB, et al. Consumers' perceptions of vape shops in Southern California: an analysis of online Yelp reviews. *Tob Induc Dis* 2014;12:22 (cited 8 April 2015). <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4258055/> <http://dx.doi.org/10.1186/s12971-014-0022-7>
- Harris JK, Moreland-Russell S, Choucair B, et al. Tweeting for and against public health policy: response to the Chicago department of public health's electronic cigarette twitter campaign. *J Med Internet Res* 2014;16:e238 (cited 8 April 2015). <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4210950/> <http://dx.doi.org/10.2196/jmir.3622>